



## Hare's Argument for Utilitarianism

Michael McDermott

*The Philosophical Quarterly*, Vol. 33, No. 133 (Oct., 1983), 386-391.

Stable URL:

<http://links.jstor.org/sici?sici=0031-8094%28198310%2933%3A133%3C386%3AHAFU%3E2.0.CO%3B2-8>

*The Philosophical Quarterly* is currently published by The Philosophical Quarterly.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/philquar.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

## DISCUSSIONS

### HARE'S ARGUMENT FOR UTILITARIANISM

BY MICHAEL McDERMOTT

Central to Hare's moral philosophy is his argument for utilitarianism. The latest and most authoritative version of the argument is in ch. 5–6 of his new book *Moral Thinking* (Oxford, 1981). Although the argument is supposed to be still that of *Freedom and Reason* 6.3 in essentials, Hare says, and I think no one will disagree, that the latest version is clearer (5.1). I will, at any rate, confine my attention to it.

Hare's argument is formulated at the "critical" level of moral thinking, and relies on the logical properties of the moral concepts used at that level. Its premises are four:

- (1) Moral judgements entail identical judgements about all cases identical in their universal properties.
- (2) "x ought to be done" entails "Let x be done".
- (3) One can assent (sincerely) to "Let x be done" iff one prefers *on balance* that x be done. (There is, on Hare's view, no logical fault in having conflicting preferences – I can prefer that x be done and at the same time prefer that x not be done. But I cannot, without logical fault, assent to both "Let x be done" and also "Let x not be done." To be able to assent to "Let x be done", it is not sufficient to have a preference that x be done – I must prefer it on balance.)
- (4) If I have full knowledge of another person's preferences, I shall myself have acquired preferences equal to his regarding what should be done to me were I in his situation.

Premise 1 asserts the universalisability of moral judgements. Premises 2 and 3 together assert the prescriptivity and overridingness (3.7, 8) of moral judgements, at the critical level. These are, in a clear enough sense, *logical* premises. Premise 4 is of a somewhat different (and somewhat mysterious) character; we may take it as saying that a man who does not have the right preferences, regarding what should be done were he in another's shoes, is making a *factual* error about what it would be like to be in those shoes. (Premise 4 is argued for in 5.3, 4; for further remarks on its logical status see 12.7, 8.)

Hare's conclusion is:

For any case where utilitarianism entails "x ought to be done",  
No one can, without logical or factual error, assent to "x ought not to be done", but  
Anyone can, without logical or factual error, assent to "x ought to be done".

This leaves open the possibility of amoralism – refusing to assent to either of the above moral judgements. (6.3, 10.7. Hare does mention in ch. 11 some prudential considerations opposed to amoralism, but at the “intuitive” level of moral thinking, not the critical; I shall not consider these.) Hare’s conclusion does, however, deny the existence of fanatics, in one sense of the term: *no one* can assent to moral judgements opposed to those of the utilitarian, without committing some kind of logical or factual error (10.4).

Hare begins (6.1) by presenting an argument to show that I cannot, without logical or factual error, assent to “I ought to do [a certain thing] to [a certain person]” if that person very much wants it not to be done to him. The argument is this: assenting to “I ought to do it to him” would commit me to “It ought to be done to me, were I in his situation”; that in turn would commit me to prescribing “Let it be done to me, were I in his situation”; but, if I fully represent to myself his situation, including his motives, I shall prescribe instead “Let it *not* be done to me, were I in his situation”.

This argument, however, raises a “problem” for Hare: the conclusion of the argument may be inconsistent with utilitarianism. If my desire to do it to him is weaker than his desire that it not be done, then in *that* case utilitarianism may deny that I ought to do it to him. But what if my desires are more intense than his? Suppose, to use Hare’s example, that “all I think I ought to do to him is move his bicycle so that I can park my car, and he has a mild aversion to my doing this”.

Here is Hare’s way out of the problem (6.2):

I can see no reason for not adopting the same solution here as we do in cases where our own preferences conflict with one another. For example, let us change the case and suppose that it is my own bicycle, and that it is moderately inconvenient to move it, but highly inconvenient not to be able to park my car; I shall then naturally move the bicycle, thinking that that is what, prudentially speaking, I ought to do, or what I want most, all in all, to do. Reverting now to the bilateral case: we have established that, if I have full knowledge of the other person’s preferences, I shall myself have acquired preferences equal to his regarding what should be done to me were I in his situation; and these are the preferences which are now conflicting with my original prescription. So we have in effect not an interpersonal conflict of preferences or prescriptions, but an intrapersonal one; both the conflicting preferences are mine. I shall therefore deal with the conflict in exactly the same way as with that between two original preferences of my own.

The solution is re-stated a paragraph later:

I am fully aware of the strength of his desire, and therefore have a desire of equal strength that, were I in his situation, the bicycle should stay where it is. But I also have my original desire to move it in order to park my car. This latter desire wins by superior strength.

It is this line of reasoning, to solve the “problem” raised by the earlier argument, which constitutes Hare’s argument for utilitarianism. Let us examine it, granting Hare’s premises.

To clarify matters, let us separate the roles of moral judge and participant: we consider what moral judgements about the case may be assented to by A, an impartial

bystander who comes to consider the case without any relevant desires of his own. C is the car-owner, B the bike-owner, and x is the moving of the bike. C prefers that x be done, B prefers that x not be done, and C's preference is stronger than B's, so utilitarianism implies "x ought to be done".

We may now reconstruct Hare's "problem" by means of two parallel arguments. When A considers what it would be like to be in C's shoes, or in B's, he acquires the following preferences (by Premise 4):

- (i) A prefers that x be done if he were in C's shoes.
- (ii) A prefers that x not be done if he were in B's shoes.

A's only desire about the case in which he occupies C's shoes is (i), so A cannot assent to "Let x not be done if A were in C's shoes" (by Premise 3). Hence (by Premise 2) A cannot assent to "x ought not to be done, if A were in C's shoes". Hence (by Premise 1) A cannot assent to "x ought not to be done (in the actual case)". Exactly parallel reasoning, appealing this time to A's preference (ii), leads to the conclusion that A cannot assent to "x ought to be done". Hare has to show that at least the second of these arguments is faulty, and that A *can* assent to "x ought to be done".

Hare's argument is that A's desires (i) and (ii) cannot be treated separately, but must be balanced against each other to see what prescriptions A can assent to. A has "conflicting" preferences, and (i) is stronger than (ii): for (i) and (ii) are equal in strength respectively to C's desire that x be done and B's desire that x not be done. So (i) "wins by superior strength": A prefers *on balance* that x be done, so he can assent to "Let x be done", and so he can assent to "x ought to be done", as implied by utilitarianism.

This solution does not work. For A's preferences (i) and (ii) are *not* conflicting. If I prefer that x be done in a certain case, and also prefer that x not be done *in that same case*, my preferences conflict, and it is their balance (which is stronger) that determines what singular prescriptions about that case I can assent to: I can assent to "Let x be done in that case" despite my preference that x not be done, if I prefer on balance that x be done, in that case. But A's preferences (i) and (ii) are *not* preferences about the same case. This is the crucial point. A's preference (i) is not a preference about the hypothetical case in which he occupies B's shoes, so it cannot be balanced against (ii) to determine what prescriptions about that hypothetical case A can assent to. What prescriptions A can assent to about a given case is determined by what he prefers on balance should be done in *that case*. If (i) and (ii) are A's only preferences, he prefers *on balance* that x not be done in the hypothetical case in which he occupies B's shoes; so he cannot assent to "Let x be done, if A were in B's shoes"; so he cannot assent to "x ought to be done, if A were in B's shoes; and so he cannot assent to "x ought to be done" – the original argument which raised the problem is valid after all.

I said earlier that Hare's argument in ch. 6 is not intended to rule out amoralism. My objection is that it rules it *in* – if A's only preferences are (i) and (i), he not only *need* not, but *cannot*, assent either to "x ought to be done" or to "x ought not to be done".

We can reformulate the objection to line up directly with Hare's formulation of his solution, in which "I" am the car-owner, and "he" is the bike-owner. Hare assumes that my "acquired preference equal to his regarding what should be done to me were I in his situation" conflicts with "my original desire" to move the bike, and so can be balanced against it. My objection is that these preferences don't conflict, since one applies to the actual case the other to a purely hypothetical case. My only preference

about the case in which I occupy his shoes is that the bike not be moved in that case, so I cannot assent to "The bike ought to be moved, if I were in his shoes".

Hare may have been led astray by the fact that my two preferences are about cases identical in their universal properties. To make the point in terms of A's preferences, (i) and (ii) are about cases identical in everything except whose shoes A occupies. But they are not numerically identical: the hypothetical case in which A occupies C's shoes is not the same hypothetical case as that in which A occupies B's shoes. Nor is there any logical or factual fault in A's having opposite preferences about cases which differ only in whose shoes A occupies, and Hare does not argue otherwise: I can assent to different *singular* prescriptions about cases which do not differ in their universal properties (though not, of course, to different moral judgements).<sup>1</sup>

To forestall another possible confusion, I should mention a quite different kind of hypothetical case A might have preferences about. Suppose A is going to be put first into B's shoes, then into C's, and x will be done either both times or not at all: which would A prefer? Well, perhaps he would prefer the former, given that in the real-world case C's preference that x be done was stronger than B's preference that x not be done (though this does not seem to be implied by Premise 4). But, whatever A's preference about such an "all-lives-seriatim" case, it is about a *different* hypothetical case to the one his preference (ii) is about. A's preference (ii) is not about an all-lives-seriatim case; it is about a case which resembles the real-world case in all its universal properties. (In the real-world case, the bike is to be moved once, or not at all – it is not part of the case that if C moves the bike – and only then – B will later become a car-owner and move C's bike.) The difficulty for Hare's argument, and his solution to it which I have criticised, centre on moral judgements, prescriptions and preferences about hypothetical cases in which the alternatives are not of the all-lives-seriatim kind.

It may be objected, finally, that my criticism of Hare depends on a misunderstanding of his account of moral judgements. I have taken prescriptivity to mean that a man's moral judgements must match his preferences for each (actual or hypothetical) case separately. An alternative account would be that a man's moral judgements about a case must match the aggregate of his preferences about all similar cases. I have taken Hare's view to be that A can assent to "x ought to be done" (in a given case) iff, in respect of *each* actual or hypothetical case with the same universal properties, A prefers on balance (the

<sup>1</sup> Could Hare escape the objection by saying that it is *not* possible (without logical or factual error) to have different preferences about cases identical in their universal properties? Not with consistency: for it is the essence of *amoralism* to say (1) "Let me do it to him", but (2) "Let it not be done to me in like circumstances" (p. 112); and Hare claims there is no logical or factual fault in amoralism. Moreover, it is hard to see any possible ground for an accusation of logical or factual error (as distinct from moral fault) in such a combination of preferences.

A possible source of confusion here might be a failure to distinguish between *cases* and *situations*. (Hare usually observes the distinction, but occasionally uses "situation" when he means "case".) A situation can be the same in different cases. For example, in our "actual" case A is just a bystander; then we ask him "How would you like it if you were in B's situation (i.e. if you were a bike-owner of similar tastes, threatened with having your bike moved)?" The question raises a *different case*, a merely hypothetical one – A is not actually a threatened bike-owner. But the situation he has to imagine himself in is the *same situation* as that B actually occupies. The identity of situations through distinct cases does not show, however, that you can't have different preferences about those cases: preferences refer not to bare situations, described without saying who is *in* them, but to cases.

balance of his preferences about *that* case) that *x* be done in that case. The alternative account is: A can assent to "x ought to be done" (in a given case) iff, in respect of *all* the actual and hypothetical cases with the same universal properties taken collectively, A prefers on balance that *x* be done in all rather than that *x* not be done in all.

I suspect that many casual readers of Hare attribute this alternative view of prescriptivity to him. But it is clearly inconsistent with the bulk of his work. There would, for instance, be no problem of "weakness of will" on the alternative view: for on that view a man might do something willingly, unequivocally and undividedly desiring to do it (in this particular case), and yet sincerely believe it was wrong.

We might ask at this point, however, must (i) and (ii) be A's *only* preferences? In the argument we have been considering, Hare attempts to show that A can, given a preference pattern consisting of (i) and (ii) only, assent to "Let *x* be done if A were in B's shoes", just as the man who *is* both bike-owner and car-owner can assent to "Let me move the bike", given his two conflicting preferences: in each case, Hare thinks, what can be assented to is determined by the relative strength of the *existing* preferences. I want now to explore a different line of thought, which, although it plays no noticeable part in ch. 6, may not be inconsistent with Hare's premises, and is perhaps suggested by brief remarks elsewhere.

Perhaps Hare could argue as follows. A must have preferences (i) and (ii), if he knows what it is like to be in C's and B's shoes. If these are his *only* preferences, he must be an amoralist. If he doesn't want to be an amoralist, he must get an *extra* preference to balance either (i) or (ii); if he wants to assent to some moral judgement, he must "universalise his prescriptions" – *change* his pattern of preferences so that he can assent to all the prescriptions about hypothetical cases to which he would then be committed by the universalisability of moral judgements. And, since (i) is stronger than (ii), the best way to universalise his prescriptions is to overcome (ii): A must somehow acquire a preference that *x* be done if A were in B's shoes, and this preference must be at least a little stronger than (ii); if he does this, A can assent to "Let *x* be done" for both hypothetical cases.

Whether this reasoning is inconsistent with Hare's premises depends on how Premise 4 is interpreted. If A's preferences regarding what should be done if he were in B's shoes must be *on balance* equal to B's – the *strong* interpretation of Premise 4 – there is no escape from amoralism. But if A's preferences need only *include* preferences equal to B's – the *weak* interpretation – there is room to manoeuvre; so let us adopt the weak interpretation of Premise 4.

But now, how exactly is the suggested Harean reasoning an argument for utilitarianism? If A starts with preferences (i) and (ii), and wants to universalise his preferences, it is claimed to be *best* to add preferences which outweigh (ii), the weaker preference. *Best* in what way? Perhaps this change constitutes a *smaller* change in his overall preference pattern than would be required to outweigh (i), the stronger preference; but why is that *better*? Is it more *prudent*, perhaps? When I am in a situation where all available courses of action involve the frustration of one or other of my preferences, it may be prudent to choose the alternative which frustrates the weaker preference; but our present concern is to compare *changes* in a preference-pattern, not degrees of preference-satisfaction. It might be relevant if we could show that change in a utilitarian direction yielded a preference pattern which enabled A to get more preference-satisfaction; but to argue in this way we would need to assume that *x*, the act required by utilitarianism, *will* be performed.

So the smaller change in A's preferences, leading him to accept utilitarianism, does not seem to be more prudent; and even if it were, that would not establish Hare's conclusion, that moral judgements contrary to utilitarianism always involve logical or factual error. Furthermore, if A *started off* with a strong preference that x be done if he were in C's shoes, (this is not inconsistent with our assumption that A starts off with no desires about the *actual* case), there would be no ground for suspecting him of imprudence, let alone logical or factual error, in assenting to the anti-utilitarian judgement "x ought not to be done". (A could then acquire the preference that x be done if he were in C's shoes, equal in strength to C's, and still prefer on balance that x not be done if he were in C's shoes.)

My conclusions, then, are:

On the strong interpretation of Premise 4, it implies (in conjunction with Hare's other premises) that acceptance of utilitarianism must involve logical or factual error.

On the weak interpretation of Premise 4, there is no reason to believe that acceptance of moral judgements contrary to utilitarianism must involve logical or factual error, or even imprudence.<sup>2</sup>

*University of Sydney*

## McTAGGART, SCHLESINGER, AND THE TWO-DIMENSIONAL TIME HYPOTHESIS

BY L. NATHAN OAKLANDER

In "The Unreality of Time" and *The Nature of Existence*, McTaggart offers a positive conception of time and then provides arguments that purport to establish that such a concept is contradictory and therefore cannot be applied to reality ([7], [8]). Although there are few, if any, philosophers who agree with McTaggart's conclusion that time is unreal, there are many who accept the general position that time involves passage or tense (cf. [2], [4], [6], [10], [11]). George Schlesinger is one philosopher sympathetic with McTaggart's positive views on time and in his recent book, *Aspects of Time*, he attempts to defend McTaggart's account of the passage of time ([14], pp. 30–33) by resuscitating a gambit suggested some time ago by C. D. Broad ([1], [3]; cf. [9]), namely, the notion that time has two dimensions. The purpose of this paper is to argue that Schlesinger does not vindicate McTaggart's positive conception of time, since the two-dimensional time hypothesis that it allegedly requires, is as beset with difficulties as the conception of temporal becoming it is supposed to render intelligible.

It will be useful for us to begin by stating the main elements of Schlesinger's interpretation of McTaggart's positive account of time since it is the view that he (hesitantly) intends to defend. According to McTaggart, the NOW is something that moves relative to the series of points that constitute time. Schlesinger continues:

<sup>2</sup> I am grateful to Professor Hare for many helpful comments and suggestions.